An On-line Wireless Attack Detection System Using Multi-layer Data Fusion

Francisco J. Aparicio-Navarro, Konstantinos G. Kyriakopoulos, David J. Parish
Department of Electronic and Electrical Engineering
Loughborough University
Loughborough, LE11 3TU, U.K.
e-mail: {elfja2, elkk, d.j.parish}@lboro.ac.uk

Abstract—**Computer networks and more specifically wireless communication networks are increasingly becoming susceptible to more sophisticated and untraceable attacks. Most of the current Intrusion Detection Systems either focus on just one layer of observation or use a limited number of metrics without proper data fusion techniques. However, the true status of a network is rarely accurately detectable by examining only one network layer. This paper describes a synergistic approach of fusing decisions of whether an attack takes place by using multiple measurements from different layers of wireless communication networks. The described method is implemented on a live system that monitors a wireless network in real time and gives an indication of whether a malicious frame exists or not. This is achieved by analysing specific metrics and comparing them against historical data. The proposed system assigns for each metric a belief of whether an attack takes place or not. The beliefs from different metrics are fused with the Dempster-Shafer technique with the ultimate goal of limiting false alarms by combining beliefs from various network layers. The on-line experimental results show that cross-layer techniques and data fusion perform more efficiently compared to conventional methods.**

Keywords-Cross-layer measurements; Data fusion; Dempster-Shafer; Wireless attacks; Wi-Fi

I. INTRODUCTION

In the last few years, computer networks and more specifically wireless communication networks are increasingly becoming susceptible to more sophisticated and untraceable attacks. The number of attacks and intrusion attempts that target wireless networks are constantly increasing. Therefore, finding a major solution for securing the wireless networks against these attacks has become a priority for researchers.

The implementation of network monitoring tools, such as Intrusion Detection Systems (IDSs), is fundamental in security infrastructures, especially in wireless networks, in order to provide another level of defence for detecting such attacks and protecting the network systems. The final goal of the IDSs is to identify malicious frames that circulate through the wireless network as actual attacks, and to not misinterpret legitimate frames as malicious. However, these tools do not prevent the attacks or intrusions from occurring.

As explained in more detail in [9], IDS are generally categorised as anomaly and misuse intrusion detection. The former tries to generate a profile of normal utilisation of the wireless resources, and raises an intrusion alarm if the analysed wireless traffic deviates from this normal profile. Anomaly IDSs are able to detect attacks not previously known. However, they can generate high rates of False Positive (FP) alarms. Alternatively, misuse IDS uses signatures of known attacks to identify attacks. This methodology generates low rates of FP, but is unable to detect novel attacks, leading to a high rate of False Negatives (FN).

Despite all the advances in developing IDSs, wireless network systems keep suffering from numerous attacks. This affirmation is based on different and miscellaneous reasons. One reason is because most of the implemented approaches are misuse IDS. As commented above, this method lacks efficiency in detecting novel attacks. Another reason is that many researchers, e.g. [10], keep evaluating their systems in an off-line environment by using the dataset from KDD'99 cup [11]. Undoubtedly that this dataset is very useful for evaluating the systems because it provides a framework in which the number and type of attacks is accurately known. However, this approach remains an in vitro process, and does not consider the profile of real data traffic of a network. Finally, and more importantly for our presented study, most of the current IDSs either focus on just one layer of observation (i.e. MAC layer) or use a limited number of metrics without proper data fusion techniques. However, the true status of a network is rarely accurately detectable by examining only one network layer.

As many researchers have previously demonstrated [3, 5], a cross-layer approach may offer a collaborative decision among layers, potentially resulting in higher detection accuracy rate with lower numbers of FN and FP. Hence, utilising a cross-layer approach may help towards automating the overall process of detecting and mitigating wireless network attacks.

In this work, we address a fundamental classification problem found in computer monitoring. This is determining whether a captured frame is of malicious intent or not. In order to leverage the knowledge from multiple measurements across multiple layers, data fusion techniques are used and specifically Dempster-Shafer (D-S). D-S has been used before in the computer monitoring field with the purpose of combining beliefs of the same metric at the same layer among nodes [1, 2].

It should be noted that the proposed methodology has been implemented in a real system running on-line and giving results in real time. The actual system consists of a computer with a wireless card on a monitoring mode that collects frames with the TShark [8] open source tool. The proposed algorithm takes as input the raw monitoring information, isolates specific

This work has been funded by the Engineering and Physical Sciences Research Council (EPSRC).

metrics from multiple layers and applies a data fusion technique that calculates the ultimate probability of whether an attack is taking place or not.

The contribution of this work is that, in contrast to the current work [1, 2, 3], the proposed methodology in this paper combines metrics from multiple layers and fuses the information with the D-S technique for a synergistic approach towards detecting attacks in wireless networks. The aim of our methodology requires the system to be of low cost and online, scalable and the concept should be applicable to other wireless technologies apart from the tested IEEE 802.11 standard.

The paper is organised as follows. An explanation of the D-S data fusion algorithm along with its advantages and disadvantages is given in section II. The methodology and testbed are presented in section III. In section IV the results obtained with the proposed methodology are discussed and compared against single or dual combination of metrics. Finally, conclusions are given in section V.

II. DEMPSTER-SHAFER THEORY

A. Mathematical Framework

Dempster-Shafer theory of evidence method is a discipline of mathematics that combines evidence of information from multiple and heterogeneous events in order to calculate the probability of occurrence of another event.

The D-S theory starts by assuming a Universe of Discourse $\Theta = \{\theta_1, \theta_2, ..., \theta_n\}$, also called a Frame of Discernment, which is a finite set of all possible mutually exclusive propositions and hypotheses about some problem domain.

With regards to this work, the frame of discernment is comprised of A = "*Attack*" and N = "*Normal*". Assuming Θ has two outcomes $\{A, N\}$, the total number of subsets of Θ, defined by the number of hypotheses that it composes, is $2^{\Theta} = \{A, N, \{A|N\}, \emptyset\}$

Each proposition (subset) from Θ is assigned a probability or a confidence interval within [0, 1], by an observer from the mass probability function m, also known as the basic probability assignment:

$$m: 2^{\Theta} \to [0,1] \qquad if \begin{cases} m(\emptyset) = 0 \\ m(A) \geq 0, \forall A \subseteq \Theta \\ \sum_{A \subseteq \Theta} m(A) = 1 \end{cases}$$

The function $m(A)$ is defined as A's basic probability number. It describes the measure of belief that is committed exactly to hypothesis A.

In order to define the confidence interval that is given to a certain event, two functions must first be defined. These are the Belief function (Bel) and the Plausibility function (Pl). The former is a belief measure of a hypothesis A, and it sums the mass value of all the non-empty subsets of A.

$$Bel(A) = \sum_{B \subseteq A} m(B) \qquad \forall A \subseteq \Theta$$

The doubt function (Dou) is given by

$$Dou(A) = Bel(\neg A) = 1 - \sum_{B \cap A = \emptyset} m(B)$$

which accounts for all evidence that rule out the given proposition represented by A.

Similarly, the Pl function takes into account all the evidence that does not rule out the given proposition. In other words, it expresses how much we should believe in A if all currently unknown facts were to support A.

$$Pl(A) = 1 - Dou(A)$$

Thus, the true belief in hypothesis A will be along the interval $[Bel(A), Pl(A)]$. However, in practice, the values of the interval could be identical and therefore the interval becomes a unique value.

The idea behind the D-S rule of combination is to fuse the belief from two different observers into one given hypothesis.

TABLE I. EXAMPLE EVENT PROBABILITIES ASSIGNED BY m_1 AND m_2

$m_2 \backslash m_1$	$\{A\}$: 0.32	$\{N\}$: 0.25	$\{A, N\}$: 0.43
$\{A\}$: 0.35	0.11	0.09	0.15
$\{N\}$: 0.1	0.03	0.025	0.04
$\{A, N\}$: 0.55	0.18	0.14	0.24

Let m_1 and m_2 be the basic probability assignments from observer 1 and 2 respectively. The cells in the above table represent the multiplication of the m_1 belief with the m_2 belief, horizontal and vertical axis, respectively.

Their orthogonal $sum, m = m_1 \oplus m_2$, is defined as

$$m(A) = \frac{\sum_{X \cap Y = A} m_1(X) * m_2(Y)}{1 - \sum_{X \cap Y = \emptyset} m_1(X) * m_2(Y)} \quad \forall A \neq \emptyset \qquad (1)$$

If the denominator of eq. (1) is equal to zero, $K = 0$, then $m_1 \oplus m_2$ does not exist and m_1 and m_2 are said to be totally or flatly contradictory.

To easily understand how to apply the D-S algorithm, a real example from our measurements is presented. The basic probabilities for an event being "*Attack*", "*Normal*", and "*Uncertain*", can be tabulated as seen in Table I.

Firstly K is calculated from eq. (1): $K = 1 - (0.03 + 0.09) = 0.88$. Similarly, $1/K = 1.136$. As described in eq. (1), for any event E the combined belief is given by:

$$m(E) = \frac{1}{K} * \sum_{X \cap Y = E} m_1(X) * m_2(Y)$$

Therefore,

$$m(A) = 1.136 * (0.11 + 0.15 + 0.18) = 0.5$$
$$m(N) = 1.136 * (0.025 + 0.04 + 0.14) = 0.233$$
$$m(A|N) = 1.136 * (0.24) = 0.272$$

According to the results, the hypothesis more likely to be true is A, with higher belief than the other hypotheses.

B. Advantages and Disadvantages

Among the different methods such as Bayesian or Principal Component Analysis, the D-S theory of evidence was chosen as the data fusion method for three clear reasons. Firstly, D-S is able to combine evidence from multiple and heterogeneous sources. Second, D-S is suitable for detecting previously unseen attacks because it does not require a priori knowledge. Finally, and more importantly, D-S method provides the ability of managing and assigning probability to ignorance, which allows tackling a large range of problems.

In contrast, Bayesian inference requires a priori knowledge and does not allow allocation of probability to ignorance but only to an event being normal or abnormal [4].

Nevertheless, there are two main drawbacks associated with the D-S algorithm. First, the high computation complexity and second the conflicting beliefs management. The computational complexity increases exponentially with the number of possible event outcomes (Θ). If there are n elements in Θ, there will be up to 2^n - 1 focal elements for the mass functions, ignoring $\emptyset$. The combination of two mass functions needs the computation of up to 2^n intersections [4].

The frame of discernment in the proposed methodology includes two elements ($n = 2$), normal and abnormal, and therefore there will be three focal elements of belief functions, 2^2 = {*Attack*, *Normal*, {*Attack* | *Normal*}, $\emptyset$}. By using only three elements in the focal elements, the fusion method requires low computational complexity.

The conflicting belief phenomenon is nicely illustrated with an example from [4]. Given three events, {*A*, *B*, *C*} and two sensors, Sensor 1 might assign $m(A) = 0.9$, $m(B) = 0.1$ and $m(C) = 0$ as beliefs in *A*, *B* and *C* respectively. Similarly, Sensor 2 might assign $m(A) = 0$, $m(B) = 0.1$ and $m(C) = 0.9$ as beliefs in *A*, *B* and *C*. Applying the D-S algorithm on these values, the rule of combination will result with a higher belief in event *B*, which is clearly wrong. In the proposed detection algorithm of this work, each event is assigned a non-zero mass function and therefore the belief conflict phenomenon is not an issue.

III. Methodology

A. Attack Description

The most common and straight forward method for an attacker to perform a Man-in-the-Middle (MitM) attack is to do first MAC spoofing, usually by performing an ARP poisoning attack (i.e. the attacker sends messages indicating that he owns a specific MAC address). This is a well known MAC layer attack. For the purposes of this work, a MitM attack between an Access Point (AP) and a client, implemented by the Airpwn tool [6], was used experimented.

Airpwn takes advantage of the Round Trip Time (RTT) that a web server takes to respond to normal webpage requests. In that lag time, it can inject its own content onto the wireless channel of an AP. If an attacker near the victim is running the Airpwn tool, it will see the legal request from the client and immediately respond with its own HTML code. Due to the fact that there are no hops between the attacker and the victim, it takes the attacker much less time to respond. When the client receives the data, it will assume the original request was answered and process the injected code.

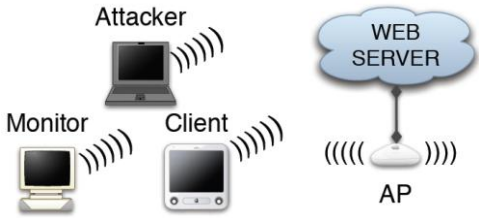

Figure 1. Testbed and steps of attack for Airpwn.

When the client receives the data, it will assume the original request was answered by the legal AP and process the injected code. Even though the attack is launched at the application layer by injecting an HTTP packet, the actual attack is practical only because there are no mechanisms in the IEEE 802.11 standard to prevent a misbehaving node from injecting their own malicious code in the form of valid WiFi frames.

Using scripts, Airpwn injects carefully crafted response code that could cause harm of varying severity. Less dangerous effects to the victim could include replacing the advert contents of a specific website with different ones; more dangerous activity could include redirecting the victim web browser to a phishing type of web site.

In our experiments, one type of attack was launched against the client. This attack code was a default option in the Airpwn suite. In this attack, the attacker listens for requests for images hosted on the web site and injects its own images. In addition, the attacker injects Reset (RST) frames in the TCP layer, making the client to request the remaining objects of the web site.

B. Testbed and Methodology

The testbed where the experiments took place can be seen in Fig. 1. It includes a client associated with an AP and accessing webpages hosted on the Internet across different geographical locations. These are China, Spain, UK and two different webpages hosted in US. For the purpose of detecting the attacks, a computer with a wireless card on monitoring mode utilised the TShark monitoring software for collecting frames. The monitoring node and the attacker were running the BackTrack Linux operating system and all the devices except from the AP used Atheros chipset in their wireless cards. The AP was a Cisco Linksys model WRT54GL.

The captured metrics of the monitoring node include all information transmitted between the AP and the client, but also include the injected malicious frames from the attacker. From all the collected information, three metrics were identified that if appropriately used could give evidence of a MitM attack. These metrics are the Received Signal Strength Indication (RSSI), the Injection Rate (or Transmission Rate), and the Time To Live (TTL) value.

Both, the TShark monitoring process and our attack detector process were concurrently launched in the computer acting as the monitor. When the wireless card listen a frame, TShark monitoring tool identifies and isolates the respective RSSI, Injection Rate and TTL of each single frame. These metrics values were the input of the attack detector process, which applied the D-S method for deciding whether the current analysed frame is legitimate or malicious.

It should be noted that the attacker was placed very close to the AP, around 1.5 meters away. This positioning of the equipment made the detection of attacks much more difficult as the RSSI values of the attacker could become identical to these of the AP. The proposed methodology can be seen as a flow chart in Fig. 2.

The attack detector process that we apply in our methodology is explained here in more detail. From the information within the captured frames, the statistical mode of RSSI, TTL and Injection Rate is calculated for a specific frame window size. In our experiments we used 20 frames for generating the mode. The metrics values of RSSI, TTL and Injection Rate from every captured frame are compared against the statistical mode of the current frames window. The beliefs for the hypotheses *Attack* and *Normal* are dependent on the distance of the value of each metric of the current frame from the calculated mode. The actual values of each belief are chosen experimentally and the intuition behind this is that the longer the distance from the mode, the higher the belief in the hypotheses *Attack* as this indicates a departure from the normality.

In order to evaluate our proposed methodology, the time instance and the number of malicious frames for each experiment are known by manually analysing the captured files. Therefore, the results of FP and FN can easily be constructed.

Our proposed methodology performs the detection in an online mode, in which the metric values of the captured frames are fed as input to the proposed algorithm for analysis, one by one, which automatically gives a belief of whether each frame is malicious or not. According to these beliefs, our proposed methodology performs the decision making in real-time, for every single frame, of whether an attack is taking place or not.

IV. Practical Results and Discussion

In this section the results from the proposed cross-layer methodology are presented and compared against single layer metrics and against the cross-layer technique using just two metrics. The results are evaluated by comparing the FN, FP and the Overall Success Rate (OSR). The OSR is the number of correct classifications divided by the total number of classifications as described in [7]. It should be noted that the performance of a technique should be judged by considering all metrics FP, FN and OSR. Considering just the OSR metric is misleading as OSR is heavily influenced by the performance of FP while mostly ignoring the performance of FN.

The cross-layer results are presented in Table II and are the best results overall and for each individual experiment. The algorithm detects all malicious frames except for some FN and FP results that occur while launching the Airpwn attack, when the client visits the UK and US_01 websites. Some FP results occur when the client visits the US_02 website, while launching the attack. These false alarms occur because the value of the used metrics, RSSI, Injection Rate and TTL, coincide with the values of the estimated mode. This happens because consecutive injected forged frames skew the value of the estimated mode away from the mode value of legitimate frames. In other words, the malicious frames outnumber the normal frames and therefore the estimated mode corresponds to the malicious frames rather than the legitimate frames.

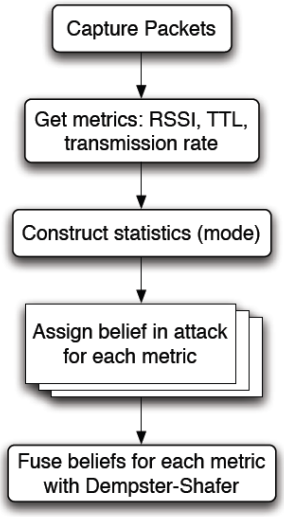

Figure 2. Methodology flowchart.

As a result, the values of all metrics derived from the attack frames are close or similar to the estimated mode, producing a high number of FN. For the same reason, the distance of all the metrics values of the legitimate frames from the mode increases. This in turn results in a higher number of FP.

In the case of double metrics RSSI and Injection Rate (Table III) the results are of poor performance with high FN percentage reaching almost 27% in two cases. The results present also a high number of FP. However, five of these FP are lower than 1%, and all of them lower than 4%.

The combination of the metrics TTL and Injection Rate, presented in Table IV, produces overall results very similar to Table III. In that case, the percentages of FN and FP are slightly higher than the previous case.

In the case of the single metric TTL (Table V) the detection accuracy clearly decreases in comparison to all the other presented results. It is noteworthy that all of the experiments generate a FN percentage higher that 3%, reaching 22% in one of the cases. The results of using the single metric Injection Rate and RSSI are not presented. However, it is worthy to say that the detection accuracy of both procedures are overall similar to the results of the single metric TTL.

These results are a clear example showcasing that the combination of beliefs from different metrics yields an improved performance.

V. Conclusions and Future Work

This paper argues that the conventional approach of using single metrics for detecting attacks in wireless networks is sometimes inefficient, inaccurate and misleading. Similarly, techniques involving multiple metrics without utilising a proper data fusion technique lack efficiency. To this aim, the authors of this work have proposed a new approach for detecting wireless network attacks, involving combining beliefs from sensors of multiple layers of observation to produce a collective decision on whether an attack is taking place or not.

The beliefs from different metrics are combined with the Dempster-Shafer theory of evidence, a mathematical framework for the representation of uncertainty, with the ultimate

goal of limiting false alarms and improving the overall performance.

TABLE II. CROSS LAYER RESULTS UTILISING RSSI, INJ. RATE AND TTL

Web Site	Type	OSR (%)	False Neg. (%)	False Pos. (%)
China	Normal	100	0	0
	Attack	100	0	0
Spain	Normal	100	0	0
	Attack	100	0	0
UK	Normal	100	0	0
	Attack	90.45	9.55	4.70
US_01	Normal	100	0	0
	Attack	85.71	14.29	3.71
US_02	Normal	100	0	0
	Attack	100	0	0.08

TABLE III. DUAL METRIC RESULTS UTILISING INJ. RATE AND RSSI

Web Site	Type	OSR (%)	False Neg. (%)	False Pos. (%)
China	Normal	100	0	0
	Attack	100	0	0.24
Spain	Normal	100	0	0.35
	Attack	73.33	26.67	3.73
UK	Normal	100	0	0.49
	Attack	73.17	26.83	2.59
US_01	Normal	100	0	0
	Attack	100	0	0.75
US_02	Normal	100	0	0.29
	Attack	91.84	8.16	1.22

TABLE IV. DUAL METRIC RESULTS UTILISING INJ. RATE AND TTL

Web Site	Type	OSR (%)	False Neg. (%)	False Pos. (%)
China	Normal	100	0	0.04
	Attack	100	0	0.43
Spain	Normal	100	0	0
	Attack	93.33	6.67	0.77
UK	Normal	100	0	1.17
	Attack	95.62	4.38	1.78
US_01	Normal	100	0	0.27
	Attack	82.35	17.65	3.85
US_02	Normal	100	0	2.33
	Attack	78.74	21.26	5.76

In this paper, the authors have demonstrated experimentally on a real wireless network that combining beliefs in real time from multiple metrics in various layers outperforms the efficiency and accuracy of single metrics. For detecting the injected attacks, the cross-layer results are the best for each individual experiment. The FN results are produced because consecutive injected forged frames skew the value of the estimated mode away from the mode value of legitimate frames. Clearly, this is a conceptual issue of window based algorithms.

TABLE V. SINGLE METRIC RESULTS UTILISING TTL

Web Site	Type	OSR (%)	False Neg. (%)	False Pos. (%)
China	Normal	100	0	4.06
	Attack	100	0	2.74
Spain	Normal	100	0	5.24
	Attack	100	0	5.22
UK	Normal	100	0	14.58
	Attack	97.50	2.50	20.73
US_01	Normal	100	0	9.60
	Attack	97.37	2.63	6.67
US_02	Normal	100	0	22.26
	Attack	87.32	12.68	12.42

As for future work, an important issue to consider is how to automate the assignment of beliefs and the adaptive selection of appropriate metrics using data mining techniques. In addition, the authors are planning to examine other types of wireless attacks.

REFERENCES

[1] V. Chatzigiannakis, G. Androulidakis, K. Pelechrrinis, S. Papavassiliou, and V. Maglaris, "Data fusion algorithms for network anomaly detection: classification and evaluation," in *Proc. ICNS 2007*, Athens, Greece. June 19-25. 2007, pp. 50–56.

[2] A. G. Fragkiadakis, V. A. Siris, and A. P. Traganitis, "Effective and robust detection of jamming attacks," in *Proc. Future Network and MobileSummit 2010*, Florence, Italy. June 16-18. 2010, pp. 1–8.

[3] G. Thamilarasu, S. Mishra, and R. Sridhar, "A cross-layer approach to detect jamming attacks in wireless ad hoc networks," in *Proc. MILCOM 2006*, Washington, District of Columbia, USA. October 23-27. 2006, pp. 1–7.

[4] Q. Chen, and U. Aickelin, "Anomaly detection using the dempster-shafer method," in *Proc. DMIN 2006*, Las Vegas, Nevada, USA. June 26-29. 2006, pp. 232–240.

[5] X. Wang, J. S. Wong, F. Stanley, and S. Basu, "Cross-layer based anomaly detection in wireless mesh networks," in *Proc. SAINT 2009*, Bellevue, Washington, USA. July 20-24. 2009, pp. 9–15.

[6] "Airpwn sourceforge website, " Available: http://airpwn.sourceforge.net/Airpwn.html.

[7] I. H. Witten, and E. Frank, *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufmann, 2nd edition, 2005.

[8] "TShark website," Available: http://www.wireshark.org/docs/man-pages/tshark.html.

[9] Y. Bai, and H. Kobayashi, "Intrusion detection systems: technology and development," in *Proc. AINA 2003*, Xi'an, China. March 27-29. 2003, pp. 710–715.

[10] K. Prakobphol, and J. Zhan, "A novel outliner detection scheme for network intrusion detection systems," in *Proc. ISA 2008*, Busan, Korea. April 24-28. 2008, pp. 555–560.

[11] "KDD cup 1999 dataset website," Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.